Subdifferential Calculus for the Value Function in Nonconvex Dynamic Optimization in Banach Spaces

著者	Sagara Nobusumi
出版者	Institute of Comparative Economic Studies, Hosei University
journal or publication title	比較経済研究所ワーキングペーパー
volume	162
page range	1-31
year	2011-02-22
URL	http://hdl.handle.net/10114/7219

Bayes 的学習過程と異時点間意思決定の経済分析：資産価格変動と異質的期待の役割
シリーズ No. 2

Subdifferential Calculus for the Value Function in Nonconvex Dynamic Optimization in Banach Spaces

Nobusumi Sagara

Subdifferential Calculus for the Value Function in Nonconvex Dynamic Optimization in Banach Spaces*

Nobusumi Sagara

Faculty of Economics, Hosei University
4342, Aihara, Machida, Tokyo, 194–0298, Japan
e-mail: nsagara@hosei.ac.jp

October 4, 2010

Abstract

This paper explores nonsmooth analysis for infinite-horizon dynamic programming in discrete time without convexity assumptions, exploiting Clarke subdifferentials for locally Lipschitz functions defined on Banach spaces. The results of subdifferential calculus for marginal functions are efficaciously employed for the Bellman equation to derive the necessary condition for optimality, and the Gâteaux and Fréchet differentiability of the value function. The main results are applied to the problem of optimal economic growth under uncertainty. Using Yosida–Hewitt decomposition of finitely additive measures, an alternative technique to obtain L^1-price systems is proposed without convexity assumptions.

Key Words: Clarke subdifferential; Gâteaux and Fréchet derivatives; Marginal function; Value function; Support price; Yosida–Hewitt decomposition.
MSC2000: Primary 49J52, 49J50; Secondary 49L20, 91B62
OR/MS subject classification: Primary: Dynamic programming/optimal control/Models; Secondary: Programming/Nondifferentiable

*This research is supported by a Grant-in-Aid for Scientific Research (No. 21530277) from the Ministry of Education, Culture, Sports, Science and Technology.

1 Introduction

In dynamic optimization theory, value functions play a crucial role in characterizing optimal programs. For instance, duality theory for convex variational problems with an infinite horizon exploits a subgradient of the value function as the support price of an optimal program (see McKenzie [24], Michel [25], Papageorgiou [28], Takekuma [36] and Weitzman [40]). When the value function is differentiable, this characterization of optimality yields an Euler equation. It has been well demonstrated that the differentiability of the value function of an infinite horizon optimization problem is guaranteed in both discrete and continuous time by a classical result of Benveniste and Scheinkman [7], under the smoothness and the convexity assumptions of the primitive data and the interiority condition of optimal programs (see also Bonnisseau and Le Van [11]).

However, for an investigation of optimal programs, especially in dynamic behavior, the differentiability of the value function is unnecessary in most cases because under smoothness assumptions, the traditional approach employing Euler equations is sufficiently powerful even without convexity assumptions that may lead to nonsmooth value functions. Moreover, the characterization of optimal programs in terms of support prices is useful only for convex models yielding concave value functions. Thus, beyond convex analysis, we need to pursue a different direction to characterize an optimal path in nonsmooth, nonconvex models that may not include support prices.

The purpose of this paper is to explore nonsmooth analysis for infinite-horizon dynamic programming in discrete time without convexity assumptions. To deal with nonsmoothness and nonconvexity together, we exploit Clarke directional derivatives and Clarke subdifferentials for locally Lipschitz functions defined on Banach spaces, tractable analytical instruments developed by Clarke [13]. The classical result of Danskin [15] and its variant by Milgrom and Segal [26] concerning the directional derivative of a marginal function of a parametric optimization problem are formulated in Banach space settings, and the results of Clarke subdifferentials of marginal functions are efficaciously employed for the Bellman equation to derive the necessary condition for optimality, and the Gâteaux and Fréchet differentiability of the value function.

The organization of the paper is as follows. Section 2 presents the terminology and preliminary results of subdifferential calculus for locally Lipschitz functions in Banach spaces. We introduce Clarke directional derivatives, Clarke subdifferentials, notions of the regularity and the Gâteaux, Fréchet and strict differentiability. A basic reference for the topic treated here is a classic study by Clarke [13], while Borwein and Zhu [12], Mordukhovich

[27], Schirotzek [34] and Zălinescu [41] are also rich sources of relevant results. In Section 3, we develop subdifferential calculus for marginal functions in parametric optimization and present some significant and new results of the Gâteaux and Fréchet derivatives of marginal functions for later use.

The main results of this paper are stated in Sections 4 and 5. Section 4 sheds light upon the simple observation that the value function of a dynamic programming problem is obtained as a marginal function of a parametric optimization problem, and applies the results in Section 3 to the subdifferential calculus for the value function. We present a necessary condition for optimality in terms of the Clarke subdifferential, which reduces in the convex case to the existence of a support price system along the lines of Gale [19], McKenzie [24], Michel [25] and Weitzman [40]. We then extend the result by Benveniste and Scheinkman [7] to the nonconvex case to obtain the Gâteaux and Fréchet differentiability of the value function. We also reproduce another important result of Amir [3], who demonstrated the differentiability of the value function under the assumption that the return function is a C^1-supermodular function defined on a vector lattice. A significant improvement of this result is proposed.

Section 5 is devoted to the application of the main results to the problem of optimal economic growth under uncertainty with L^∞-commodity spaces along the lines of Dana [14], Evstigneev [18], Jeanjean [21], Papageorgiou [28], Radner [31], Takekuma [36] and Zilcha [42, 43]. We show the existence of an L^1-price system for an optimal program by the effective use of the celebrated Yosida–Hewitt decomposition theorem of finitely additive measures (see Yosida and Hewitt [39]). An alternative technique to obtain L^1-price systems is proposed without convexity assumptions, where support price systems are unavailable.

2 Preliminaries

2.1 Derivatives and Subdifferentials

Let X be a real Banach space with the dual system $\langle X, X^* \rangle$, where X^* is the norm dual of X and endowed with the weak*-topology. The norm of a generic element x in X is denoted by $\|x\|$ and a generic element in X^* is denoted by p. The *domain* of an extended real-valued function $f : X \to \mathbb{R} \cup \{\pm\infty\}$ is the set $\operatorname{dom} f = \{x \in X \mid f(x) < +\infty\}$. The function f is *Lipschitz of rank* $K \geq 0$ near a given point $x \in X$ at which f is finite if there exists a neighborhood U of x such that $|f(y) - f(z)| \leq K\|y - z\|$ for every $y, z \in U$. The function f is said to be *locally Lipschitz* on an open set $U \subset X$ if f is

Lipschitz near x for every $x \in U$.

The following result reinforces the well-known property of the locally Lipschitz continuity of convex functions that are locally bounded above (see Borwein and Zhu [12, Theorem 4.1.3]).

Proposition 2.1. *If $f : X \to \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous and convex, then it is locally Lipschitz on* $\operatorname{int}(\operatorname{dom} f)$.

The (*one-sided*) *directional derivative* of a function $f : U \to \mathbb{R}$ at $x \in U$ in the direction $h \in X$ is defined by:

$$f'(x; h) = \lim_{\lambda \downarrow 0} \frac{f(x + \lambda h) - f(x)}{\lambda}$$

when this limit exists in $\mathbb{R}$. We say that f is *Gâteaux differentiable* at x provided that there exists an element $D_G f(x) \in X^*$, called the *Gâteaux derivative* at x, such that $f'(x; h) = \langle D_G f(x), h \rangle$ for every $h \in X$. If $f : X \to \mathbb{R}$ is continuous and convex on a separable Banach space X, then the set of points at which f is Gâteaux differentiable is a dense G_δ set in X.

The *Clarke directional derivative* of $f : U \to \mathbb{R}$ at x in the direction h, denoted by $f^\circ(x; h)$, is defined as follows:

$$f^\circ(x; h) = \limsup_{\substack{y \to x \\ \lambda \downarrow 0}} \frac{f(y + \lambda h) - f(y)}{\lambda}.$$

The function f is *regular* at x provided that f admits the directional derivative $f'(x; h)$ at x with $f'(x; h) = f^\circ(x; h)$ for every $h \in X$. The *Clarke subdifferential* of f at x, denoted by $\partial f(x)$, is defined by:

$$\partial f(x) = \{p \in X^* \mid \langle p, h \rangle \le f^\circ(x; h) \ \forall h \in X\}.$$

Because $\partial f(\cdot)$ gives rise to a set-valued mapping from U into X^*, we denote it by $\partial f : U \rightrightarrows X^*$. It is known that if f is Lipschitz near x, then $\partial f(x)$ is nonempty, convex and weakly* compact. Moreover, the Clarke directional derivative is the support function of the Clarke subdifferential in the sense that:

$$f^\circ(x; h) = \max\{\langle p, h \rangle \mid p \in \partial f(x)\} \quad \text{for every } h \in X. \tag{2.1}$$

If $f : X \to \mathbb{R}$ is Lipschitz near $x \in X$, and regular and Gâteaux differentiable at x, then $\partial f(x) = \{D_G f(x)\}$.

If $f : U \to \mathbb{R}$ is a convex function on an open convex set U and is Lipschitz near $x \in U$, then it is regular at x, and its Clarke subdifferential at x reduces to its *subdifferential* at x in the sense of convex analysis:

$$\partial f(x) = \{p \in X^* \mid \langle p, h \rangle \le f'(x; h) \ \forall h \in X\}. \tag{2.2}$$

If $f : X \to \mathbb{R}$ is continuous and convex, then it is Gâteaux differentiable at $x \in \operatorname{dom} f$ if and only if $\partial f(x)$ is a singleton.

We say that $f : U \to \mathbb{R}$ is *Fréchet differentiable* at $x \in U$, and $D_F f(x) \in X^*$ is the *Fréchet derivative* at x provided that:

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - \langle D_F f(x), h \rangle}{\|h\|} = 0.$$

We say that f is *continuously differentiable* at x if it has Fréchet derivative $D_F f(x)$, and $y \mapsto D_F f(y) \in X^*$ is (norm-to-norm) continuous on the neighborhood of x. If $f : X \to \mathbb{R}$ is continuous and convex on X and X^* is separable, then the set of points at which f is Fréchet differentiable is a dense G_δ set in X. When $X = \mathbb{R}^n$, Gâteaux and Fréchet differentiability coincide for locally Lipschitz functions.

For later use, we cite a useful criterion of the Fréchet differentiability for Gâteaux-differentiable convex functions defined on Banach spaces (see Zălinescu [41, Corollary 3.3.6]).

Proposition 2.2. *Let $f : X \to \mathbb{R}$ be continuous, convex and Gâteaux differentiable on the neighborhood U of x. Then f is Fréchet differentiable at x if and only if $D_G f : U \to X^*$ is continuous at x.*

Proof of the following result can be found in Schirotzek [34, Propositions 3.4.2 and 7.3.9; Remark 7.4.2].

Proposition 2.3. *If $f : U \to \mathbb{R}$ is continuously differentiable at $x \in U$, then f is regular at x, locally Lipschitz and $\partial f(x) = \{D_F f(x)\}$.*

We say that $f : U \to \mathbb{R}$ is *strictly* (*Hadamard*) *differentiable* at $x \in U$ and $D_s f(x) \in X^*$ is the *strict* (*Hadamard*) *derivative* at x provided that:

$$\lim_{\substack{h \to 0 \\ y \to x}} \frac{f(y+h) - f(y) - \langle D_s f(x), h \rangle}{\|h\|} = 0.$$

Strict differentiability implies Gâteaux differentiability. Moreover, if $f : U \to \mathbb{R}$ is strictly differentiable at $x \in U$, then it is regular at x. When $X = \mathbb{R}^n$, strict differentiability also implies Fréchet differentiability.

3 Subdifferential Calculus for Parametric Optimization

3.1 The Clarke Subdifferential of the Marginal Function

We investigate the differentiability properties of the marginal function of a parametric optimization problem of the form $\max_{y\in Y} f(x,y)$, where $f : X \times Y \to \mathbb{R}$ is an upper semicontinuous function, where X is a Banach space and Y is a compact topological space. We denote this by $f'_x(x,y;h)$ and $f^\circ_x(x,y;h)$, the directional and the Clarke directional derivatives of $f(\,\cdot\,,y)$ at $x \in Y$ in the direction $h \in X$, respectively, and by $\partial_x f(x,y)$ the Clarke subdifferential of $f(\,\cdot\,,y)$ at x. Similarly, $D_{G,x}f(x,y)$ and $D_{F,x}f(x,y)$ denote the Gâteaux and the Fréchet derivatives of $f(\,\cdot\,,y)$ at x, respectively.

We present the following result, which is a generalization of Sagara [33] for the case where $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$.

Lemma 3.1 (Sagara). *Let U be an open subset of X and Y be a topological space. If $f : U \times Y \to \mathbb{R}$ is upper semicontinuous such that $f(\,\cdot\,,y)$ is Lipschitz for every $y \in Y$, then $\partial_x f : U \times Y \rightrightarrows X^*$ is upper semicontinuous and weakly* compact-valued.*

Proof. Denote the open unit ball in X by B. Arbitrarily take $\varepsilon > 0$ and $x_0 \in U$ satisfying $x_0 + \varepsilon B \subset U$, choose any $0 < \varepsilon' < \varepsilon$ and put $\eta = \varepsilon - \varepsilon'$. Then, $\{x + \eta B \mid x \in x_0 + \varepsilon' B\}$ is an open covering of $x_0 + \varepsilon B$. We claim that the function $g_\eta(\,\cdot\,,\cdot\,;h) : (x_0 + (\varepsilon-\eta)B) \times Y \to \mathbb{R}$ defined by:

$$g_\eta(x,y;h) = \sup_{\substack{x' \in x+\eta B \\ \lambda \in (0,\eta)}} \frac{f(x'+\lambda h, y) - f(x',y)}{\lambda}$$

is upper semicontinuous for every $h \in X$. To this end, note that the supremum is finite because of the Lipschitz continuity of $f(\,\cdot\,,y)$.

Define the set-valued mapping $\Gamma_\eta : (x_0 + (\varepsilon-\eta)B) \times Y \rightrightarrows X \times Y \times \mathbb{R}$ by $\Gamma_\eta(x,y) = (x+\eta B) \times \{y\} \times (0,\eta)$. Then, $g_\eta(\,\cdot\,,\cdot\,;h)$ is the marginal function of the continuous function:

$$(x',y',\lambda) \mapsto \frac{f(x'+\lambda h, y') - f(x',y')}{\lambda}$$

on $(x_0 + (\varepsilon-\eta)B) \times Y \times (0,\eta)$ maximized over $(x',y',\lambda) \in \Gamma_\eta(x,y)$. Because Γ_η is continuous, applying the maximum theorem (see Dontchev and Zolezzi [16, Theorem IX.1.5]) yields the continuity of $g_\eta(\,\cdot\,,\cdot\,;h)$ on $(x_0 + \varepsilon B) \times Y$

for every $\eta \in (0, \varepsilon)$. This implies that $\inf_{\eta \in (0,\varepsilon)} g_\eta(x, y; h)$ is the pointwise infimum of the continuous functions $\{g_\eta(\,\cdot\,, \cdot\,; h)\}$, but the infimum coincides with $f_x^\circ(x, y; h)$ by definition. Therefore, $f_x^\circ(\,\cdot\,, \cdot\,; h)$ is upper semicontinuous on $(x_0 + \varepsilon B) \times Y$ for every $h \in X$.

Let $(x^\nu, y^\nu) \to (x, y)$ and $p^\nu \xrightarrow{w^*} p$ be convergent nets in $(x_0+\varepsilon B)\times Y$ and X^*, respectively, satisfying $p^\nu \in \partial_x f(x^\nu, y^\nu)$ for each ν. By (2.1), we have $\langle p^\nu, h\rangle \le f^\circ(x^\nu, y^\nu, h)$ for each ν and $h \in X$. Thus, taking the lim sup on both sides of this inequality yields $\langle p, h\rangle = \lim_\nu \langle p^\nu, h\rangle \le \limsup_\nu f^\circ(x^\nu, y^\nu, h) \le f^\circ(x, y, h)$ for every $h \in X$ by the upper semicontinuity of $f_x^\circ(\,\cdot\,, \cdot\,; h)$. This shows that $p \in \partial_x f(x, y)$ and the weakly* compact-valued subdifferential mapping $\partial_x f$ has the closed graph and hence is upper semicontinuous on $(x_0 + \varepsilon B) \times Y$ (see Aliprantis and Border [1, Corollary 17.17]). Because the choice of $\varepsilon > 0$ and $x_0 \in U$ is arbitrary insomuch $x_0 + \varepsilon B \subset U$, the upper semicontinuity of $\partial_x f$ on $U \times Y$ immediately follows. □

Proposition 3.1 (Clarke). *Let U be an open subset of X, Y a compact metric space and $f : U \times Y \to \mathbb{R}$ an upper semicontinuous function such that $f(\,\cdot\,, y)$ is Lipschitz of rank K for every $y \in Y$. Define the marginal function $m : U \to \mathbb{R}$ as:*

$$m(x) = \max\{f(x, y) \mid y \in Y\}$$

and the solution mapping $M : U \rightrightarrows Y$ as:

$$M(x) = \{y \in Y \mid f(x, y) = m(x)\}.$$

Then, m is Lipschitz of rank K and the following formula holds:

$$\partial m(x) \subset \left\{ \int_Y \partial_x f(x, y) d\mu(y) \mid \mu \in \mathscr{P}(M(x)) \right\} \tag{3.1}$$

for every $x \in U$, where $\mathscr{P}(M(x))$ denotes the family of probability Radon measures on Y supported on $M(x)$. If, moreover, $f(\,\cdot\,, y)$ is regular at x for every $y \in M(x)$, then m is regular at x and the equality holds for (3.1).

The intended meaning of the inclusion (3.1) is as follows. Every element $p \in \partial m(x)$ is written as $\langle p, h\rangle = \int \langle p(y), h\rangle d\mu(y)$ for every $h \in X$ with some $\mu \in \mathscr{P}(M(x))$, where $p : Y \to X^*$ is a measurable selection of $\partial f(x, \cdot) : Y \rightrightarrows X^*$ such that $p(y) \in \partial_x f(x, y)$ μ-a.e. $y \in Y$. (Here, X^* is endowed with the Borel σ-algebra generated by the weak* topology.)

Proof. Because $\partial_x f : U \times Y \rightrightarrows X^*$ is upper semicontinuous, convex and weakly*-compact-valued by Lemma 3.1, it is *closed* at every point $(x, y) \in U \times Y$ in the sense of Clarke [13, Definition 2.8.1]. Thus, the theorem reduces to the special case of [13, Theorem 2.8.2]. □

Proposition 3.1 is very powerful. Formulas for the directional, Gâteaux and Fréchet derivatives of the marginal function can be derived systematically from it, as demonstrated in Subsections 3.2 and 3.3. Formula (3.1) is an elaboration of Pshenichnyi [30, Theorem 3.5]. When the Lipschitz condition in Proposition 3.1 is replaced with the convexity of $f(\,\cdot\,, y)$, inclusion (3.1) becomes equality, which is the formula established by Ioffe and Levin [20], because the convexity guarantees that $f(\,\cdot\,, y)$ is regular at x (see Clarke [13, Propositions 2.2.6 and 2.3.6]).

Theorem 3.1. *Let Y be a compact subset of a Banach space. If $f : X \times Y \to \mathbb{R}$ is locally Lipschitz and regular at $(x, y) \in X \times Y$ with $y \in M(x) \cap \operatorname{int} Y$, then:*

$$\{p \in X^* \mid (p, 0) \in \partial f(x, y)\} \subset \partial m(x). \tag{3.2}$$

Proof. It follows from $y \in M(x) \cap \operatorname{int} Y$ that for an arbitrarily given direction $k \in Y$, we have $f(x, y + \lambda k) \leq f(x, y)$ for all sufficiently small $\lambda > 0$. This implies that $f_y^\circ(x, y; k) = f_y'(x, y; k) \leq 0$ for every $k \in X$ because of the regularity of f. Thus, if $q \in \partial_y f(x, y)$, then equations (2.1) and (2.2) yields $\langle q, k\rangle \leq f_y'(x, y; k) \leq 0$ for every $k \in Y$, and, hence, $q = 0$. Therefore, $\partial_y f(x, y) = \{0\}$. Choose any $(p, q) \in \partial f(x, y)$. Because the regularity of f implies that $\partial f(x, y) \subset \partial_x f(x, y) \times \partial_y f(x, y)$ (see Clarke [13, Proposition 2.3.15]), we have $(p, q) \in \partial_x f(x, y) \times \partial_y f(x, y)$ and $q = 0$. Thus, $(p, 0) \in \partial f(x, y)$. Formula (3.1) with equality implies that $\partial_x f(x, y) \subset \partial m(x)$ because $\mathscr{P}(M(x))$ includes a unit mass of each point in $M(x)$. Therefore, $p \in \partial m(x)$. $\square$

When f is upper semicontinuous and concave, inclusion (3.2) becomes equality with no additional condition. Indeed, one can dispense with Lipschitz continuity, compactness, regularity and interiority in Theorem 3.1 (see Aubin and Ekeland [5, Proposition 4.5.6]). Formula (3.2) has been intensively improved without compactness assumptions. For example, Penot [29] exploits the Fréchet subdifferential and Thibault [37] uses the Mordukhovich subdifferential (the sequential limits of the Fréchet subdifferential) for the subdifferential calculus of marginal functions. For a systematic treatment of marginal functions, see Mordukhovich [27, Subsections 1.3.4 and 3.2.1].

3.2 Directional Derivative of the Marginal Function

The advantage of Proposition 3.1 is exemplified by the observation that it reduces to the classical result of Danskin [15, Theorem III. I], who formulated the directional derivative of the marginal function in the case where the objective function has the partial derivative regarding parameter values in a finite dimensional Euclidean space.

Theorem 3.2 (Danskin). *Let U be an open subset of a Banach space X with its separable dual and Y be a compact metric space. If $f : U \times Y \to \mathbb{R}$ is an upper semicontinuous function such that $f(\,\cdot\,, y)$ is Lipschitz of rank K for every $y \in Y$ and $f(\,\cdot\,, y)$ is regular at x for every $y \in M(x)$, then m is regular at x and admits a directional derivative satisfying:*

$$m'(x;h) = \max_{y \in M(x)} f'_x(x,y;h) \quad \textit{for every } h \in X. \tag{3.3}$$

Proof. Denote by $p(\cdot) : Y \to X^*$ a measurable selection of $\partial_x f(x,\cdot\,) : Y \rightrightarrows X^*$, which exists by Proposition 3.1. Let $x, h \in X$ be given arbitrarily. It is easy to verify the inequality:

$$\max_{p(\cdot) \in \partial_x f(x,\cdot\,)} \int_Y \langle p(y), h \rangle d\mu(y) \leq \int_Y \Big[\max_{p \in \partial_x f(x,y)} \langle p, h \rangle \Big] d\mu(y).$$

Define the set-valued mapping $Q : Y \rightrightarrows X^*$ by:

$$Q(y) = \left\{ q \in \partial_x f(x,y) \mid \langle q, h \rangle = \max_{p \in \partial_x f(x,y)} \langle p, h \rangle \right\}.$$

Because $\partial_x f(x,\cdot\,)$ is upper semicontinuous and has weakly* compact values by Lemma 3.1, and X^* is separable in the norm topology, by the measurable maximum theorem (see Aliprantis and Border [1, Theorem 18.19]), one can choose a measurable selection $q(\cdot) : Y \to X^*$ of Q. Because $q(\cdot)$ is also a measurable selection of $\partial_x f(x,\cdot\,)$, we have:

$$\int_Y \Big[\max_{p \in \partial_x f(x,y)} \langle p, h \rangle \Big] d\mu(y) = \int_Y \langle q(y), h \rangle d\mu(y) \leq \max_{p(\cdot) \in \partial_x f(x,\cdot\,)} \int_Y \langle p(y), h \rangle d\mu(y),$$

and, hence:

$$\int_Y \Big[\max_{p \in \partial_x f(x,y)} \langle p, h \rangle \Big] d\mu(y) = \max_{p(\cdot) \in \partial_x f(x,\cdot\,)} \int_Y \langle p(y), h \rangle d\mu(y) \tag{3.4}$$

for every $\mu \in \mathscr{P}(M(x))$. Because m is regular at x by Proposition 3.1, the equality in (3.1) yields:

$$\begin{aligned} m'(x;h) = \max_{p \in \partial m(x)} \langle p, h \rangle &= \max_{\substack{\mu \in \mathscr{P}(M(x)) \\ p(\cdot) \in \partial_x f(x,\cdot)}} \int_Y \langle p(y), h \rangle \, d\mu(y) \\ &= \max_{\mu \in \mathscr{P}(M(x))} \int_Y \Big[\max_{p \in \partial_x f(x,y)} \langle p, h \rangle \Big] d\mu(y) \end{aligned}$$

$$
\begin{aligned}
&= \max_{\mu \in \mathscr{P}(M(x))} \int_Y f_x^\circ(x, y; h) d\mu(y) \\
&= \max_{\mu \in \mathscr{P}(M(x))} \int_Y f_x'(x, y; h) d\mu(y) \\
&= \max_{y \in M(x)} f_x'(x, y; h),
\end{aligned}
$$

where the second equality in the first line is a consequence of (2.1) and (2.2), the second line uses (3.4), the third line employs (2.1), the fourth line exploits the regularity of f and the last line follows from the observation that $\mathscr{P}(M(x))$ contains a unit mass of each point in $M(x)$. □

Theorem 3.3. *Let U be an open subset of X and Y be a compact metric space. If $f : U \times Y \to \mathbb{R}$ is an upper semicontinuous function such that $f(\,\cdot\,, y)$ is Lipschitz of rank K for every $y \in Y$, $f(\,\cdot\,, y)$ is regular at x for every $y \in M(x)$ and the set $\{f_x'(x, y; h) \in \mathbb{R} \mid y \in M(x)\}$ is a singleton for every $h \in X$, then:*

$$\partial m(x) = \partial_x f(x, y) \quad \textit{for every } y \in M(x).$$

Proof. Because $f(\,\cdot\,, y)$ is regular at x for every $y \in M(x)$, we have $f_x'(x, y; h) = f_x^\circ(x, y; h) = \max_{p \in \partial_x f(x,y)} \langle p, h \rangle$ for every $h \in X$. This means that for every $y, y' \in M(x)$, the support functions of weakly*-compact convex sets $\partial_x f(x, y)$ and $\partial_x f(x, y')$ are identical; hence, $\max_{p \in \partial_x f(x,y)} \langle p, h \rangle = \max_{p \in \partial_x f(x,y')} \langle p, h \rangle$ for every $h \in X$. From the Hörmander theorem (see Clarke [13, Theorem 2.1.4]; Schirotzek [34, Theorem 2.3.1]), it follows that $\partial_x f(x, y) = \partial_x f(x, y')$ for every $y, y' \in M(x)$. Because formula (3.1) holds with equality, we obtain $\partial m(x) = \partial_x f(x, y)$ for every $y \in M(x)$ because $\mathscr{P}(M(x))$ contains a unit mass of each point in $M(x)$. □

There are two important cases where it is possible to verify by Theorem 3.3 the single-valuedness of the directional derivatives of $f_x'(x, y; h)$ with respect to $y \in M(x)$; one is the case where f is a concave function on a convex set (Theorem 3.4) and the other where f is a supermodular function on a vector lattice (Theorem 3.5).

Theorem 3.4. *Let Y be a convex subset of a vector space. Suppose that $M(x)$ is nonempty at $x \in X$. If $f : X \times Y \to \mathbb{R}$ is concave and $f(\,\cdot\,, y)$ has a two-sided directional derivative at x for every $y \in M(x)$, then the set $\{f_x'(x, y; h) \in \mathbb{R} \mid y \in M(x)\}$ is a singleton for every $h \in X$.*

Proof. By symmetry, it suffices to show that $f_x'(x, y_0; h) \geq f_x'(x, y_1; h)$ for every $y_0, y_1 \in M(x)$ and $h \in X$. Suppose on the contrary that the points y_0

and y_1 exist in $M(x)$ such that $f'_x(x, y_0; h) < f'_x(x, y_1; h)$ for some nonzero $h \in X$. Because of the equality:

$$-f'_x(x, y_0; h) = \lim_{\lambda \downarrow 0} \frac{f(x - \lambda h, y_0) - f(x, y_0)}{\lambda}$$

following from the definition of the (two-sided) directional derivative, we have $-f(x - \lambda h, y_0) + f(x, y_0) < f(x + \lambda h, y_1) - f(x, y_1)$ for all sufficiently small $\lambda > 0$. This inequality yields:

$$\begin{aligned} m(x) = \frac{1}{2}(f(x, y_0) + f(x, y_1)) &< \frac{1}{2}(f(x - \lambda h, y_0) + f(x + \lambda h, y_1)) \\ &\leq f(x, \tfrac{1}{2}(y_0 + y_1)) = m(x), \end{aligned}$$

a contradiction in view of $\frac{1}{2}(y_0 + y_1) \in M(x)$ from the concavity of f. $\square$

Parametric Lattice Programming

Let X be a Banach space that is also a vector lattice. Because we do not necessarily assume here that X is a Banach lattice, the lattice operations in X may not be norm continuous; but this poses no problem with the limiting argument. A function $f : X \to \mathbb{R} \cup \{-\infty\}$ is *supermodular* if $f(x) + f(x') \leq f(x \vee x') + f(x \wedge x')$ for every $x, x' \in X$. If f is supermodular, then the set of points at which f attains its maximum is a sublattice of X (see Topkis [38]).

Theorem 3.5. *Let X be a Banach space that is also a vector lattice and Y be a lattice. Suppose that $M(x)$ is nonempty at $x \in X$. If $f : X \times Y \to \mathbb{R}$ is supermodular, and $f(\,\cdot\,, y)$ is upper semicontinuous and has a two-sided directional derivative at x for every $y \in M(x)$, then the set $\{f'_x(x, y; h) \in \mathbb{R} \mid y \in M(x)\}$ is a singleton for every $h \in X$.*

Proof. Take any $y_0, y_1 \in M(x)$. If $f'_x(x, y_0; h) < f'_x(x, y_1; h)$ for some nonzero $h \in X$, then from the proof of Theorem 3.4, it follows that there exists some $\varepsilon > 0$ such that $f(x, y_0) + f(x, y_1) < f(x - \lambda h, y_0) + f(x + \lambda h, y_1) - \lambda\varepsilon$ for all sufficiently small $\lambda > 0$. Because $(x + \lambda h) \vee (x - \lambda h) = x + \lambda|h|$ and $(x + \lambda h) \wedge (x - \lambda h) = x - \lambda|h|$, the supermodularity of f implies that $f(x - \lambda h, y_0) + f(x + \lambda h, y_1) \leq f(x + \lambda|h|, y_0 \vee y_1) + f(x - \lambda|h|, y_0 \wedge y_1)$. We thus have:

$$f(x, y_0) + f(x, y_1) < f(x + \lambda|h|, y_0 \vee y_1) + f(x + \lambda|h|, y_0 \wedge y_1) - \lambda\varepsilon \quad (3.5)$$

for all sufficiently small $\lambda > 0$. Note that $y_0 \vee y_1$ and $y_0 \wedge y_1$ belong to $M(x)$ because $M(x)$ is a sublattice of X. By the upper semicontinuity of $f(\,\cdot\,, y)$ at

x for every $y \in M(x)$, there exists a sequence $\lambda^\nu \downarrow 0$ such that $f(x+\lambda^\nu|h|, y_0 \vee y_1) < f(x, y_0 \vee y_1) + \frac{\lambda^\nu \varepsilon}{2}$ and $f(x - \lambda^\nu|h|, y_0 \wedge y_1) < f(x, y_0 \wedge y_1) + \frac{\lambda^\nu \varepsilon}{2}$ for each $\nu = 1, 2, \ldots$. Then by (3.5), we have $2m(x) = f(x, y_0) + f(x, y_1) < f(x, y_0 \vee y_1) + f(x, y_0 \wedge y_1) = 2m(x)$, a contradiction. □

3.3 Differentiability of the Marginal Function

Another consequence of Proposition 3.1 revealed with the aid of the Danskin formula is the Gâteaux and Fréchet differentiability of the marginal function.

Theorem 3.6. *Let U be an open subset of Banach space X with its separable dual, Y be a compact metric space and $f : U \times Y \to \mathbb{R}$ be an upper semicontinuous function such that $f(\,\cdot\,, y)$ is Lipschitz of rank K for every $y \in Y$.*

(i) *If $f(\,\cdot\,, y)$ is regular and Gâteaux differentiable at $x \in U$ for every $y \in M(x)$, then m admits a Gâteaux derivative at x if and only if $\{D_{G,x} f(x, y) \in X^* \mid y \in M(x)\}$ is a singleton. The Gâteaux derivative of m satisfies the formula:*

$$D_G m(x) = D_{G,x} f(x, y) \quad \text{for every } y \in M(x).$$

(ii) *If $f(\,\cdot\,, y)$ is continuously differentiable at $x \in U$ for every $y \in M(x)$, then m is continuously differentiable at x if and only if $\{D_{F,x} f(x, y) \in X^* \mid y \in M(x)\}$ is a singleton. The Fréchet derivative of m satisfies the formula:*

$$D_F m(x) = D_{F,x} f(x, y) \quad \text{for every } y \in M(x).$$

Proof. (i): If the set $\{D_{G,x} f(x, y) \in X^* \mid y \in M(x)\}$ is a singleton, then $m'(x; h) = \langle D_{G,x} f(x, y), h\rangle$ for every $y \in M(x)$ and $h \in X$ in view of the Danskin formula (3.3). This means that $m'(x; \cdot\,) \in X^*$, and, hence, $m'(x; \cdot\,) = D_G m(x) = D_{G,x} f(x, y)$ for every $y \in M(x)$. Conversely, if m admits a Gâteaux derivative at x, then the Danskin formula implies that $\langle D_G m(x), -h\rangle = -\min_{y \in M(x)} \langle D_{G,x} f(x, y), h\rangle$. We thus have:

$$\langle D_G m(x), h\rangle = \min_{y \in M(x)} \langle D_{G,x} f(x, y), h\rangle = \max_{y \in M(x)} \langle D_{G,x} f(x, y), h\rangle$$

for every $h \in X$. Therefore, $\{D_{G,x} f(x, y) \in X^* \mid y \in M(x)\}$ must be a singleton and $D_G m(x) = D_{G,x} f(x, y)$ for every $y \in M(x)$.

(ii): Because $f(\,\cdot\,, y)$ is regular at x for every $y \in M(x)$ by Proposition 2.3, the rest of the proof is the same as for (i). The continuity of the Fréchet derivative $D_F m$ at x follows from that of $D_F f(\,\cdot\,, y)$ at x. □

Lemma 3.2. *Let Y be a compact subset of a Banach space. If $f : X \times Y \to \mathbb{R}$ is locally Lipschitz, then for every $x \in X$ there exists a constant K for which $f(\,\cdot\,,y)$ is Lipschitz of rank K near x for every $y \in Y$.*

Proof. Because f is locally Lipschitz, for every $(x, y) \in X \times \operatorname{int} Y$, there exists a neighborhood $U_y \times V_y$ of (x, y) on which f is Lipschitz of rank K_y. Because Y is compact, there exist points $y_1, \ldots, y_n$ in Y for which $V_{y_1}, \ldots, V_{y_n}$ constitute a finite open covering of Y, and f is Lipschitz of rank K_{y_i} on each $U_{y_i} \times V_{y_i}$, $i = 1, \ldots, n$. Put $K = \max\{K_{y_1}, \ldots, K_{y_n}\}$. By construction, $f(\,\cdot\,, y)$ is Lipschitz of rank K on $\bigcap_{i=1}^n U_{y_i}$ for every $y \in Y$. □

Corollary 3.1. *Let X be a Banach space with its separable dual, let Y be a compact convex subset of a Banach space and $f : X \times Y \to \mathbb{R}$ be upper semicontinuous and concave.*

(i) *If $f(\,\cdot\,, y)$ is regular and Gâteaux differentiable at $x \in X$ for every $y \in M(x)$, then m admits a Gâteaux derivative at x and:*

$$D_G m(x) = D_{G,x} f(x, y) \quad \textit{for every } y \in M(x).$$

(ii) *If $f(\,\cdot\,, y)$ is continuously differentiable $x \in X$ for every $y \in M(x)$, then m is continuously differentiable $x \in X$ and:*

$$D_F m(x) = D_{F,x} f(x, y) \quad \textit{for every } y \in M(x).$$

Proof. (i): Because f is locally Lipschitz by Proposition 2.1, $f(\,\cdot\,, y)$ is Lipschitz of rank K on a convex neighborhood U of x for every $y \in Y$ by virtue of Lemma 3.2. The Gâteaux differentiability of m and the formula follows from Theorems 3.4 and 3.6.

(ii): Because $f(\,\cdot\,, y)$ is regular at x for every $y \in M(x)$ by Proposition 2.3, m has a Gâteaux derivative with $D_G m(x) = D_{F,x} f(x, y)$ for every $y \in M(x)$, by (i). Therefore, $D_G m(x) = D_F m(x)$ by Proposition 2.2. □

Corollary 3.2. *Let X be a Banach space with its separable dual and a vector lattice, Y be a compact lattice, and $f : X \times Y \to \mathbb{R}$ be an upper semicontinuous supermodular function such that $f(\,\cdot\,, y)$ is Lipschitz of rank K for every $y \in Y$.*

(i) *If $f(\,\cdot\,, y)$ is regular and Gâteaux differentiable at $x \in X$ for every $y \in M(x)$, then m admits a Gâteaux derivative at x and:*

$$D_G m(x) = D_{G,x} f(x, y) \quad \textit{for every } y \in M(x).$$

(ii) *If $f(\,\cdot\,, y)$ is continuously differentiable at $x \in X$ for every $y \in M(x)$, then m is continuously differentiable $x \in X$ and:*

$$D_F m(x) = D_{F,x} f(x, y) \quad \textit{for every } y \in M(x).$$

Proof. The formula follows from Theorems 3.5 and 3.6. □

4 Dynamic Programming with an Infinite Horizon

4.1 Description of the Model

Let T be a set of nonnegative integers, the set of time horizons indexed by $t = 0, 1, \dots$. For each $t \in T$, let X_t be a nonempty subset of Banach space X. A function $f_t : X_t \times X_{t+1} \to \mathbb{R}$ is a return function, and a set-valued mapping $\Gamma_t : X_t \rightrightarrows X_{t+1}$ describes a feasibility constraint. The graph of Γ_t is denoted by $G(\Gamma_t)$.

An *admissible program* starting from period $t \in T$ with an initial condition $x \in X_t$ is an element $(x_t, x_{t+1}, \dots)$ in the product space $\prod_{s=t}^{\infty} X_s$ satisfying $x_{s+1} \in \Gamma_s(x_s)$ for every $s \in T$ and $x_t = x$. The set of admissible programs from t with $x_t = x$ is denoted by $\mathscr{A}_t(x)$, which gives rise to a set-valued mapping $\mathscr{A}_t : X_t \rightrightarrows \prod_{s=t}^{\infty} X_s$. An admissible program $(x_0, x_1, \dots) \in \mathscr{A}_0(x)$ for a given $x \in X_0$ is *optimal* if it maximizes the objective function $\sum_{t \in T} f_t$ over $\mathscr{A}_0(x)$.

For the primitive $\{X_t, \Gamma_t, f_t \mid t \in T\}$ of the model, the following basic assumptions are in force, which are quite standard in dynamic programming with finite dimensional program spaces (see Stokey and Lucas [35, Chapter 4]).

Assumption 4.1. (i) X_t is nonempty and closed.

(ii) $\Gamma_t : X_t \rightrightarrows X_{t+1}$ is continuous and has nonempty, compact values.

(iii) $f_t : X_t \times X_{t+1} \to \mathbb{R}$ is continuous.

(iv) $\displaystyle\sum_{t \in T} \sup_{(x,y) \in G(\Gamma_t)} |f_t(x, y)| < +\infty$.

Define the *value function* $v_t : X_t \to \mathbb{R}$, which is defined by:

$$v_t(x) = \max \left\{ \sum_{s=t}^{\infty} f_s(x_s, x_{s+1}) \mid (x_t, x_{t+1}, \dots) \in \mathscr{A}_t(x) \right\}.$$

It follows from the *Bellman principle of optimality* that every optimal program $(x_0, x_1, \dots) \in \mathscr{A}_0(x_0)$ satisfies:

$$v_t(x_t) = f_t(x_t, x_{t+1}) + v_{t+1}(x_{t+1}) \quad \text{for every } t \in T. \tag{BP}$$

It is verified by way of standard argument that the value function satisfies the *Bellman equation*:

$$v_t(x) = \max_{y \in \Gamma_t(x)} \{f_t(x, y) + v_{t+1}(y)\} \quad \text{for every } x \in X_t, \tag{BE}$$

which demonstrates that the value function is obtained as a marginal function investigated in Section 3. This simple observation motivates the introduction of solution mapping $M_t : X_t \rightrightarrows X_{t+1}$ for (BE) by:

$$M_t(x) = \{y \in \Gamma_t(x) \mid v_t(x) = f_t(x, y) + v_{t+1}(y)\}.$$

By (BP) and (BE), an admissible program $(x_0, x_1, \dots) \in \mathscr{A}_0(x)$ is optimal if and only if $x_{t+1} \in M_t(x_t)$ for every $t \in T$. Assumption 4.1 is sufficient to ensure that there exists an optimal program, v_t is continuous and has values in $\mathbb{R}$, and M_t is upper semicontinuous and compact-valued.

4.2 Auxiliary Results

Throughout this paper, we retain Assumption 4.1 as a basic hypothesis. The conditions for the interiority of optimal programs and the Lipschitz conditions on f_t are required in the sequel.

Assumption 4.2. (i) $\operatorname{int} X_t$ is nonempty.

(ii) For every $x \in X_t$ and $y \in M_t(x)$: $(x, y) \in \operatorname{int} G(\Gamma_t)$.

The Lipschitz continuity of the return function is a point where we depart from the standard assumption, which replaces the traditional smoothness or convexity hypotheses. Two alternative Lipschitz continuity conditions on f_t are employed in the sequel.

Assumption 4.3. $f_t(\,\cdot\,, y)$ is locally Lipschitz uniformly in $y \in X_{t+1}$.

Assumption 4.3*. $f_t : X_t \times X_{t+1} \to \mathbb{R}$ is locally Lipschitz.

The precise meaning of Assumption 4.3 is that for every $x \in X_t$ there exists a neighborhood U_t of x and $K_t \geq 0$ such that f_t is Lipschitz of rank K_t on U_t for every $y \in X_{t+1}$. While the Lipschitz bound K_t may depend on x, it is constant on U_t.

The next result, because of Askri and Le Van [4], is very useful for our purpose, and the proof of this is somewhat simplified here.

Lemma 4.1 (Askri and Le Van). *If Assumptions 4.1 and 4.2 are satisfied, then for every $x \in X_t$, there exists a neighborhood U_t of x and a compact subset Y_{t+1} of X_{t+1} such that $M_t(x') \subset Y_{t+1} \subset \Gamma_t(x')$ for every $x' \in U_t$.*

Proof. By Assumption 4.2(ii), for every $y \in M_t(x)$, there exists a neighborhood $U_y \times V_y \subset X_t \times X_{t+1}$ of (x, y) such that $(x', y') \in \operatorname{int} G(\Gamma_t)$ for every $(x', y') \in U_y \times V_y$. Because $M_t(x)$ is compact by Assumption 4.1(ii) and (iii), there exist points $y^1, \dots, y^n$ in $M_t(x)$ such that $V_{y^1}, \dots, V_{y^n}$ constitutes a finite open covering of $M_t(x)$. Put $U_t' = \bigcap_{i=1}^n U_{y^i}$ and $V_{t+1} = \bigcup_{i=1}^n V_{y^i}$. By construction, we have $M_t(x) \subset V_{t+1} \subset \operatorname{int} \Gamma_t(x)$ and $V_{t+1} \subset \operatorname{int} \Gamma_t(x')$ for every $x' \in U_t'$. Because M_t is upper semicontinuous at $x \in X_t$, there exists a neighborhood U_t'' of x satisfying $M_t(x') \subset V_{t+1}$ for every $x' \in U_t''$. Putting $U_t = U_t' \cap U_t''$ and $Y_{t+1} = \operatorname{cl} V_{t+1}$ yields the desired inclusion because Y_{t+1} is compact by Assumption 4.1(ii). □

Lemma 4.2. *If Assumptions 4.1, 4.2 and 4.3 are satisfied, then $v_t : X_t \to \mathbb{R}$ is locally Lipschitz.*

Proof. By Lemma 4.1, there exists a neighborhood U_t of an arbitrarily given point in X_t and a compact set Y_{t+1} such that $M_t(x) \subset Y_{t+1} \subset \Gamma_t(x)$ for every $x \in U_t$. Then the Bellman equation (BE) becomes:

$$v_t(x) = \sup_{y \in Y_{t+1}} \{f_t(x, y) + v_{t+1}(y)\} \quad \text{for every } x \in U_t. \tag{4.1}$$

Because $f_t(\,\cdot\,, y)$ is locally Lipschitz uniformly in $y \in X_{t+1}$ by Assumption 4.3, we may assume without loss of generality that $|f_t(x, y) - f_t(x', y)| \le K_t\|x - x'\|$ for every $x, x' \in U_t$ and $y \in X_{t+1}$. Let $x, x' \in U_t$ be given arbitrarily. Take any $\varepsilon > 0$ and let $y \in Y_{t+1}$ be such that $v_t(x) < f_t(x, y) + v_{t+1}(y) + \varepsilon$. In view of (4.1), we have $v_t(x') \ge f_t(x', y) + v_{t+1}(y)$. We thus obtain:

$$v_t(x) - v_t(x') < f_t(x, y) - f_t(x', y) + \varepsilon \le K_t\|x - x'\| + \varepsilon.$$

Because ε is arbitrary, by interchanging the role of x and x', we obtain $|v_t(x) - v_t(x')| \le K_t\|x - x'\|$ for every $x, x' \in U_t$. Therefore, v_t is locally Lipschitz. □

Theorem 4.1. *Suppose that Assumptions 4.1, 4.2 and 4.3 are satisfied. Then:*

$$\partial v_t(x) \subset \left\{ \int_Y \partial_x f_t(x, y) d\mu(y) \mid \mu \in \mathscr{P}(M_t(x)) \right\}$$

for every $x \in X_t$. If $f_t(\,\cdot\,, y)$ is regular at x for every $y \in M_t(x)$, then v_t is regular at x and:

$$\partial_x f_t(x, y) \subset \partial v_t(x) \quad \textit{for every } y \in M_t(x).$$

If, moreover, X has its separable dual, then:

$$v_t'(x;h) = \max_{y \in M_t(x)} f_{t,x}'(x,y;h) \quad \text{for every } h \in X.$$

Proof. Choose any $x \in X_t$. By Lemma 4.1, there exists a neighborhood U_t of x and a compact set Y_{t+1} such that $M_t(x') \subset Y_{t+1} \subset \Gamma_t(x')$ for every $x' \in U_t$ and (4.1) holds. By Lemma 4.2, v_t is Lipschitz on U_t. Define $g_t : U_t \times Y_{t+1} \to \mathbb{R}$ by $g_t(x', y) = f_t(x', y) + v_{t+1}(y)$. Then, by Assumption 4.3, $g_t(\,\cdot\,, y)$ is Lipschitz of rank K_t on U_t for every $y \in Y_{t+1}$. A direct application of Proposition 3.1 and Theorem 3.2 to $U = U_t$, $Y = Y_{t+1}$, $f = g_t$ and $m = v_t$ yields the first inclusion. Because the equality holds for this inclusion under the regularity hypothesis of f_t and $\mathscr{P}(M_t(x))$ contains a unit mass of each point in $M_t(x)$, the second inclusion immediately holds. The remaining equality is a consequence of (3.3). $\square$

Assumption 4.2(ii) is somewhat stronger than the standard interiority condition such that for every $x \in X_t$ and $y \in M_t(x)$: $y \in \operatorname{int} \Gamma_t(x)$ (see Benveniste and Scheinkman [7], and Stokey and Lucas [35]). For the mitigation of this condition to ensure the smoothness of the value function under the convexity hypothesis, consult Aliprantis et al. [2], and Rincón-Zapatero and Santos [32].

4.3 Necessary Condition for Optimality

While optimality is described by the Euler equation for the smooth case, the *generalized Euler equation* (or *Euler inclusion*):

$$0 \in \partial_y f_t(x_t, x_{t+1}) + \partial_x f_{t+1}(x_{t+1}, x_{t+2}) \quad \text{for every } t \in T \tag{GE}$$

along an optimal program can be derived for the nonsmooth case from the following theorem.

Theorem 4.2. *Suppose that Assumptions 4.1, 4.2 and 4.3* are satisfied, and let $(x_0, x_1, \dots) \in \mathscr{A}_0(x_0)$ be an optimal program. If f_t is regular at (x_t, x_{t+1}), then there exists a sequence $\{p_t \in X^* \mid t \in T\}$ such that:*

(i) $p_t \in \partial v_t(x_t)$ *for every* $t \in T$.

(ii) $(p_t, -p_{t+1}) \in \partial f_t(x_t, x_{t+1})$ *for every* $t \in T$.

If, moreover, the Lipschitz bound K_t of $f_t(\,\cdot\,, x_{t+1})$ near x_t satisfies $\lim_t K_t = 0$*, then:*

(iii) $\lim_{t \to \infty} p_t = 0$.

Proof. Note first that $f_t(\,\cdot\,, y)$ is Lipschitz of rank K_t near x_t for every $y \in M_t(x_t)$ by Lemma 3.2, and v_{t+1} is Lipschitz near x_{t+1} and regular at x_{t+1} by Lemma 4.2 and Theorem 4.1. The set inclusion:

$$\{p \in X^* \mid (p, 0) \in \partial(f_t(x_t, x_{t+1}) + \partial v_{t+1}(x_{t+1}))\} \subset \partial v_t(x_t), \tag{4.2}$$

follows from Theorem 3.1. Let $(p_t, 0) \in \partial(f_t(x_t, x_{t+1}) + \partial v_{t+1}(x_{t+1}))$. Because the right-hand side of this inclusion is contained in $\partial f_t(x_t, x_{t+1}) + \{0\} \times \partial v_{t+1}(x_{t+1})$, there exists some $p_{t+1} \in \partial v_{t+1}(x_{t+1})$ such that $(p_t, -p_{t+1}) \in \partial f_t(x_t, x_{t+1})$. The inclusion of $p_t \in \partial v_t(x_t)$ follows from (4.2). Thus, assertions (i) and (ii) hold. Because the adjoint inclusion (ii) guarantees that $p_t \in \partial_x f(x_t, x_{t+1})$, we have $\|p_t\|_{X^*} \leq K_t$ (see Clarke [13, Proposition 2.1.2]), and, hence, assertion (iii) holds. □

Condition (iii) of Theorem 4.2 is the *transversality condition* at infinity. The condition on the Lipschitz bound, for instance, is typically satisfied for the discounted case where $f_t = \rho^t f$ with $\rho \in (0, 1)$ and $f : X \times X \to \mathbb{R}$ has a Fréchet (resp. Gâteaux or strict) derivative $D_{F,x} f(x, x_{t+1})$ at every point x in a neighborhood U_t of x_t on which $D_{F,x} f(\,\cdot\,, x_{t+1})$ is bounded in X^*. Then, the Lipschitz bound of $f_t(\,\cdot\,, x_{t+1})$ near x_t is given by $K_t = \rho^t \sup_{x \in U_t} \|D_{F,x} f(x, x_{t+1})\|_{X^*}$.

Support Price Theorem

Support prices play a crucial role in establishing the asymptotic behavior (turnpike property) of optimal programs (see Gale [19] and McKenzie [24]).

Definition 4.1. A sequence $\{p_t \in X^* \mid t \in T\}$ is a *support price system* for an admissible program $(x_0, x_1, \dots) \in \mathscr{A}_0(x_0)$ if the following conditions are satisfied.

(i) $v_t(x) - \langle p_t, x\rangle \leq v_t(x_t) - \langle p_t, x_t\rangle$ for every $x \in X_t$ and $t \in T$.

(ii) $f_t(x, y) + \langle p_{t+1}, y\rangle - \langle p_t, x\rangle \leq f_t(x_t, x_{t+1}) + \langle p_{t+1}, x_{t+1}\rangle - \langle p_t, x_t\rangle$ for every $(x, y) \in G(\Gamma_t)$ and $t \in T$.

Assumption 4.1 and the following convexity hypothesis imply that f_t is locally Lipschitz on $\operatorname{int}(X_t \times X_{t+1})$ by Proposition 2.1 and v_t is concave.

Assumption 4.4. (i) X_t and $G(\Gamma_t)$ are convex.

(ii) $f_t : X_t \times X_{t+1} \to \mathbb{R}$ is concave.

A support price theorem is provided by Gale [19], McKenzie [24] and Weitzman [40], and proven using the separation theorem for convex sets. Under Assumption 4.4, the necessary condition in Theorem 4.2 reduces to the supergradient inequalities for the concave functions v_t and f_t, which are evidently equivalent to the existence of a support price system.

Corollary 4.1 (Support price theorem). *If Assumptions 4.1, 4.2 and 4.4 are satisfied, then an optimal program admits a support price system.*

It is well known that under the convexity hypothesis with additional transversality conditions at infinity, the existence of a support price system is a sufficient condition for optimality (see McKenzie [24], Michel [25] and Weitzman [40]).

4.4 Differentiability of the Value Function

There is a big gap between the regularity and differentiability of functions at a given point. For the Clarke directional derivative $f^\circ_{t,x}(x,y;h)$ to be a Gâteaux derivative, it needs to be linear and continuous in h. A curious implication of the regularity of the return function at an interior optimal solution is the Gâteaux and strict differentiability of the value function.

Theorem 4.3. *Suppose that Assumptions 4.1, 4.2 and 4.3* are satisfied and X has its separable dual. Let $(x_0, x_1, \dots) \in \mathscr{A}_0(x_0)$ be an optimal program. If f_t is regular at (x_t, x_{t+1}) for every $t \in T$, then:*

(i) *$f_t(\,\cdot\,, x_{t+1})$ and v_t are Gâteaux differentiable at x_t and:*

$$D_G v_t(x_t) = D_{G,x} f_t(x_t, x_{t+1}) \quad \textit{for every } t \in T \setminus \{0\}.$$

(ii) *$f_t(x_t, \cdot\,)$ and v_{t+1} are strictly differentiable at x_{t+1} and:*

$$D_s v_{t+1}(x_{t+1}) = -D_{s,y} f_t(x_t, x_{t+1}) \quad \textit{for every } t \in T \setminus \{0\}.$$

Proof. To show the strict differentiability of v_{t+1}, notice that $x_{t+1} \in \operatorname{int} \Gamma_t(x_t)$ by Assumption 4.2(ii), and given an arbitrary direction $k \in X$, we have $f_t(x_t, x_{t+1} + \lambda k) + v_{t+1}(x_{t+1} + \lambda k) \le f_t(x_t, x_{t+1}) + v_{t+1}(x_{t+1})$ for all sufficiently small $\lambda > 0$, in view of (BP) and (BE). This implies that $f^\circ_{t,y}(x_t, x_{t+1}; k) + v^\circ_{t+1}(x_{t+1}; k) = f'_{t,y}(x_t, x_{t+1}; k) + v'_{t+1}(x_{t+1}; k) \le 0$ for every $k \in X$, because v_{t+1} is regular at x_{t+1} by Theorem 4.2. Choose any $q \in \partial_y f_t(x_t, x_{t+1})$ and $r \in \partial v_{t+1}(x_{t+1})$. Then, by (2.1) and (2.2), we have:

$$\langle q, k \rangle + \langle r, k \rangle \le f'_{t,y}(x_t, x_{t+1}; k) + v'_{t+1}(x_{t+1}; k) \le 0$$

for every $k \in X$. This implies that $q = -r$, and hence, $-\partial v_{t+1}(x_{t+1}) = \partial_y f_t(x_t, x_{t+1}) = \{q\}$. Therefore, $f_t(x_t, \cdot\,)$ and v_{t+1} are strictly differentiable at x_{t+1} and $q = D_{s,y} f_t(x_t, x_{t+1}) = -D_s v_{t+1}(x_{t+1})$ (see Clarke [13, Proposition 2.2.4]). This supports assertion (ii). Because strict differentiability implies Gâteaux differentiability, $D_s v_t(x_t) = D_G v_t(x_t) = D_{G,x} f_t(x_t, x_{t+1})$ by Theorem 3.6(i). This establishes assertion (i). □

Theorem 4.3 is striking because the Gâteaux differentiability of v_t is guaranteed by the regularity of f_t even without the differentiability of f_t. It leads to the Euler equation:

$$D_{G,y} f_t(x_t, x_{t+1}) + D_{G,x} f_t(x_{t+1}, x_{t+2}) = 0 \quad \text{for every } t \in T \setminus \{0\} \tag{E}$$

for an optimal program $(x_0, x_1, \dots) \in \mathscr{A}(x_0)$. However, it says nothing about the Gâteaux differentiability of v_0. Furthermore, it is silent about the Fréchet differentiability of the value function. To make the differentiability property complete, we introduce the following condition.

Assumption 4.5. At any $x \in X_t$, the set $\{f'_{t,x}(x, y; h) \in \mathbb{R} \mid y \in M_t(x)\}$ is a singleton for every $h \in X$.

Indeed, Assumption 4.5 is much milder than requiring the single-valuedness of the solution mapping M_t. If M_t is single-valued, then it gives rise to a *policy function* from X_t into X_{t+1}, which is guaranteed in most applications when one assumes the strict concavity of $f_t(x, \cdot\,)$ with the convexity hypothesis (Assumption 4.4). To derive the differentiability of the value function, Kim [22] introduced in the finite dimensional setting the single-valuedness of the partial derivative of f_t with respect to x evaluated on $M_t(x)$, which is rather stronger than Assumption 4.5.

Theorem 4.4 (Envelope theorem). *Suppose that Assumptions 4.1, 4.2, 4.3 and 4.5 are satisfied. If $f_t(\,\cdot\,, y)$ is regular at $x \in X_t$ for every $y \in M_t(x)$, then:*

$$\partial v_t(x) = \partial_x f_t(x, y) \quad \textit{for every } y \in M_t(x).$$

Proof. The equality is an immediate consequence of Theorems 3.3 and 4.1. □

Theorem 4.5. *Suppose that Assumptions 4.1, 4.2, 4.3* and 4.5 are satisfied and X has its separable dual.*

(i) *If $f_t(\,\cdot\,, y)$ is regular and Gâteaux differentiable at $x \in X_t$ for every $y \in M_t(x)$, then v_t is Gâteaux differentiable at x and:*

$$D_G v_t(x) = D_{G,x} f_t(x, y) \quad \textit{for every } y \in M_t(x).$$

(ii) *If* $f_t(\,\cdot\,, y)$ *is continuously differentiable at* $x \in X_t$ *for every* $y \in M_t(x)$, *then* v_t *is continuously differentiable at* x *and:*

$$D_F v_t(x) = D_{F,x} f_t(x, y) \quad \textit{for every } y \in M_t(x).$$

Proof. Assumption 4.5 implies that $f'_{t,x}(x, y; \cdot\,) = D_{G,x} f_t(x, y)$ for every $y \in M_t(x)$ for case (i) and $f'_{t,x}(x, y; \cdot\,) = D_{F,x} f_t(x, y)$ for every $y \in M_t(x)$ for case (ii). Applying Theorem 3.6 yields the result. □

The standard convexity hypothesis introduced in Assumption 4.4 takes the place of the Lipschitz condition in Assumption 4.3 and the single-valuedness of the directional derivatives in Assumption 4.5 by Theorem 3.4. Therefore, Theorem 4.5 extends the classical result of the differentiability of the value function in finite dimensional spaces by Benveniste and Scheinkman [7] to Banach spaces.

Corollary 4.2 (Benveniste and Scheinkman). *Suppose that Assumptions 4.1 and 4.4 are satisfied and* X *has a separable dual.*

(i) *If* $f_t(\,\cdot\,, y)$ *is regular and Gâteaux differentiable at* $x \in X_t$ *for every* $y \in M_t(x)$, *then* v_t *is Gâteaux differentiable at* x *and:*

$$D_G v_t(x) = D_{G,x} f_t(x, y) \quad \textit{for every } y \in M_t(x).$$

(ii) *If* $f_t(\,\cdot\,, y)$ *is continuously differentiable* $x \in X_t$ *for every* $y \in M_t(X_t)$, *then* v_t *is continuously differentiable at* x *and:*

$$D_F v_t(x) = D_{F,x} f_t(x, y) \quad \textit{for every } y \in M_t(x).$$

Lattice Programming

An alternative assumption that guarantees the single-valuedness of the directional derivatives is the supermodularity of f_t. The following lattice structure together with Assumption 4.1 is sufficient to guarantee the supermodularity of v_t.

Assumption 4.6. (i) X is a Banach space that is also a vector lattice, and X_t is a sublattice of X.

(ii) $G(\Gamma_t)$ is a sublattice of $X \times X$.

(iii) $f_t : X_t \times X_{t+1} \to \mathbb{R}$ is supermodular.

Lemma 4.3. *If Assumptions 4.1 and 4.6 are satisfied, then* $v_t : X_t \to \mathbb{R}$ *is supermodular.*

Proof. Take any $x, x' \in X_t$. Let $(x_t, x_{t+1}, \dots) \in \mathscr{A}_t(x)$ and $(x'_t, x'_{t+1}, \dots) \in \mathscr{A}_t(x')$ be optimal programs from t with initial conditions x and x', respectively. We then have:

$$\begin{aligned} v_t(x) + v_t(x') &= \sum_{s=t}^{\infty} f_s(x_s, x_{s+1}) + \sum_{s=t}^{\infty} f_s(x'_s, x'_{s+1}) \\ &\leq \sum_{s=t}^{\infty} f_s(x_s \vee x'_s, x_{s+1} \vee x'_{s+1}) + \sum_{s=t}^{\infty} f_s(x_s \wedge x'_s, x_{s+1} \wedge x'_{s+1}) \\ &\leq v_t(x \vee x') + v_t(x \wedge x'), \end{aligned}$$

where the second inequality employs the supermodularity of f_s by Assumption 4.6(iii) and the final inequality uses $x_{s+1} \vee x'_{s+1} \in \Gamma_s(x_s \vee x'_s)$ and $x_{s+1} \wedge x'_{s+1} \in \Gamma_s(x_s \wedge x'_s)$ for each $s = t, t+1, \dots$ resulting from Assumption 4.6(ii) together with the definition of the value function. □

Corollary 4.3 (Amir). *Suppose that Assumptions 4.1, 4.2, 4.3* and 4.6 are satisfied and X has a separable dual.*

(i) *If $f_t(\,\cdot\,, y)$ is regular and Gâteaux differentiable at $x \in X_t$ for every $y \in M_t(x)$, then v_t is Gâteaux differentiable at x and:*

$$D_G v_t(x) = D_{G,x} f_t(x, y) \quad \textit{for every } y \in M_t(x).$$

(ii) *If $f_t(\,\cdot\,, y)$ is continuously differentiable at $x \in X_t$ for every $y \in M_t(x)$, then v_t is continuously differentiable at x and:*

$$D_F v_t(x) = D_{F,x} f_t(x, y) \quad \textit{for every } y \in M_t(x).$$

Proof. Because $(x, y) \mapsto f_t(x, y) + v_{t+1}(y)$ is supermodular by Lemma 4.3, Assumption 4.5 holds by Theorem 3.5. The result follows from Theorem 4.5. □

5 Optimal Economic Growth under Uncertainty

5.1 Stochastic Dynamic Programming

Let $(\Omega, \mathscr{F}, P)$ be a complete probability space, where Ω is a sample space, $\mathscr{F}$ is a σ-algebra of subsets of Ω, and P is a complete probability measure on $\mathscr{F}$. By $\{\mathscr{F}_t \mid t \in T\}$, we denote a filtration by which each $\mathscr{F}_t$ is a complete sub-σ-

algebra of $\mathscr{F}$ having $\bigvee_{t\in T}\mathscr{F}_t = \mathscr{F}$. We denote by $L^\infty(\Omega, \mathscr{F}_t, P)$ the space of $\mathbb{R}^n$-valued, essentially bounded, $\mathscr{F}_t$-measurable functions on Ω, endowed with the supremum norm $\|\cdot\|_\infty$.

The primitive data of an economy is described by random technology correspondence $Z_t : \mathbb{R}^n_+ \times \Omega \rightrightarrows \mathbb{R}^n_+$ with a graph given by:

$$G(Z_t) = \{(z, z', \omega) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+ \times \Omega \mid z' \in Z_t(z, \omega)\}$$

and a random utility function $u_t : \mathbb{R}^n_+ \times \mathbb{R}^n_+ \times \Omega \to \mathbb{R}$. Given an initial capital stock $x_0 \in L^\infty(\Omega, \mathscr{F}_0, P)$, a stochastic process $\{x_t \in L^\infty(\Omega, \mathscr{F}_t, P) \mid t \in T\}$ satisfying $x_{t+1}(\omega) \in Z_{t+1}(x_t(\omega), \omega)$ a.e. $\omega \in \Omega$ for every $t \in T$ is a *feasible program*, over which the expected utility sum:

$$\sum_{t\in T} \int_\Omega u_{t+1}(x_t(\omega), x_{t+1}(\omega), \omega)dP(\omega),$$

is maximized. This is a general reduced model of capital accumulation under uncertainty investigated by Dana [14], Evstigneev [18], Jeanjean [21], Papageorgiou [28], Radner [31], Takekuma [36] and Zilcha [42, 43].

To transform the problem into dynamic programming, define the set-valued mapping $\Gamma_t : L^\infty(\Omega, \mathscr{F}_t, P) \rightrightarrows L^\infty(\Omega, \mathscr{F}_{t+1}, P)$ by:

$$\Gamma_t(x) = \{y \in L^\infty(\Omega, \mathscr{F}_{t+1}, P) \mid y(\omega) \in Z_{t+1}(x(\omega), \omega) \text{ a.e. } \omega \in \Omega\},$$

and the return function $f_t : L^\infty(\Omega, \mathscr{F}_t, P) \times L^\infty(\Omega, \mathscr{F}_{t+1}, P) \to \mathbb{R}$ by:

$$f_t(x, y) = \int_\Omega u_{t+1}(x(\omega), y(\omega), \omega)dP(\omega).$$

Then the problem turns into the Banach space setting investigated in Section 4, in which $X = L^\infty(\Omega, \mathscr{F}, P)$ and $X_t = L^\infty(\Omega, \mathscr{F}_t, P)$.

5.2 Assumptions

Compactness of the Feasible Set

Assumption 5.1. (i) $Z_t(\,\cdot\,, \omega) : \mathbb{R}^n_+ \rightrightarrows \mathbb{R}^n_+$ is continuous and has nonempty values a.e. $\omega \in \Omega$.

(ii) $G(Z_t)$ is $\mathscr{B}(\mathbb{R}^n) \times \mathscr{B}(\mathbb{R}^n) \times \mathscr{F}_t$-measurable.

(iii) There exists an $\mathscr{F}_t$-measurable, bounded function $\hat{\sigma}_t : \Omega \to \mathbb{R}$ such that $\|z'\| \le \hat{\sigma}_t(\omega)$ for every $(z, z', \omega) \in G(Z_t)$.

Let $Z_t^x : \Omega \rightrightarrows \mathbb{R}^n_+$ be the set-valued mapping defined by $Z_t^x(\omega) := Z_t(x(\omega), \omega)$ for $x \in L^\infty(\Omega, \mathscr{F}_t, P)$, and $\mathscr{S}(Z_t^x, \mathscr{F}_{t+1})$ be the set of $\mathscr{F}_{t+1}$-measurable selections of Z_t^x.

Lemma 5.1. *If Assumption 5.1 is satisfied, then $\mathscr{S}(Z_t^x, \mathscr{F}_{t+1})$ is nonempty, bounded, closed and $\Gamma_t(x) = \mathscr{S}(Z_t^x, \mathscr{F}_{t+1})$ for every $x \in L^\infty(\Omega, \mathscr{F}_t, P)$.*

Proof. Under Assumption 5.1(ii), by Aumann's measurable selection theorem (see Aliprantis and Border [1, Theorem 18.26]; Aumann [6]; Clarke [13, Theorem 4.1.1]), there exists a measurable selection of Z_t^x, i.e., $\mathscr{F}_t$-measurable function $\sigma_t : \Omega \to \mathbb{R}^n$ such that $\sigma_t(\omega) \in Z_t^x(\omega)$ for every $\omega \in \Omega$. From $\mathscr{F}_t \subset \mathscr{F}_{t+1}$, it follows that σ_t is $\mathscr{F}_{t+1}$-measurable, and, hence, $\mathscr{S}(Z_t^x, \mathscr{F}_{t+1})$ is nonempty for every $x \in L^\infty(\Omega, \mathscr{F}_t, P)$. Therefore, $\mathscr{S}(Z_t^x, \mathscr{F}_{t+1}) \subset \Gamma_t(x)$. The converse inclusion $\Gamma_t(x) \subset \mathscr{S}(Z_t^x, \mathscr{F}_{t+1})$ is evident. By Assumption 5.1(iii), we have $\|\sigma\|_\infty \leq \|\hat{\sigma}_t\|_\infty < +\infty$ for every $\sigma \in \mathscr{S}(Z_t^x, \mathscr{F}_{t+1})$. Thus, $\Gamma_t(x)$ is bounded. The closedness of $\Gamma_t(x)$ follows from Assumption 5.1(i) $\square$

Let Π_t be the set of all finite collections $\pi = \{A_1, \dots, A_m\}$ of disjoint sets in $\mathscr{F}_t$ with positive measure. Define the order of Π_t by $\pi \leq \pi'$ to mean that each set in π is, except for a set of measure zero, a union of sets in π', which makes Π_t a directed set. For every $\pi = \{A_1, \dots, A_m\} \in \Pi_t$ and $f \in L^\infty(\Omega, \mathscr{F}_t, P)$, define the function f^π by the following equations:

$$f^\pi(\omega) = \begin{cases} \dfrac{1}{P(A_i)} \displaystyle\int_{A_i} f(\omega) dP(\omega) & \text{if } \omega \in A_i, \\ 0 & \text{if } \omega \notin \bigcup_{i=1}^n A_i. \end{cases}$$

Then, $\|f^\pi - f\|_\infty \to 0$ for every $f \in L^\infty(\Omega, \mathscr{F}_t, P)$ (see Dunford and Schwartz [17, Proof of Theorem IV.8.18]).

Weak*-compact sets in L^∞ can be obtained easily because boundedness captures relative weak*compactness. On the contrary, (norm) compact sets in L^∞ are characterized by the following "quite general criterion ..., which, however, is somewhat difficult to apply to specific cases ([17, p. 297])".

Proposition 5.1 (M. Riesz). *A bounded subset $\mathscr{K}$ of $L^\infty(\Omega, \mathscr{F}_t, P)$ is relatively compact if and only if:*

$$\lim_{\pi \in \Pi_t} \sup_{f \in \mathscr{K}} \|f^\pi - f\|_\infty = 0.$$

Applying Proposition 5.1 to $\mathscr{S}(Z_t^x, \mathscr{F}_{t+1})$ in place of $\mathscr{K}$ in the probability space $(\Omega, \mathscr{F}_{t+1}, P)$ yields the compactness of $\mathscr{S}(Z_t^x, \mathscr{F}_{t+1})$, which in turn is equivalent to the compactness of $\Gamma_t(x)$ by Lemma 5.1.

Assumption 5.2. For every $x \in L^\infty(\Omega, \mathscr{F}_t, P)$:

$$\lim_{\pi \in \Pi_{t+1}} \sup_{\sigma \in \mathscr{S}(Z_t^x, \mathscr{F}_{t+1})} \|\sigma^\pi - \sigma\|_\infty = 0.$$

Under Assumptions 5.1 and 5.2, it is evident that $\Gamma_t : L^\infty(\Omega, \mathscr{F}_t, P) \rightrightarrows L^\infty(\Omega, \mathscr{F}_{t+1}, P)$ is continuous and has nonempty compact values.

Lipschitz Continuity and Regularity of the Return Function

Assumption 5.3. (i) $u_t(z, z', \cdot) : \Omega \to \mathbb{R}$ is $\mathscr{F}_t$-measurable for every $(z, z') \in \mathbb{R}^n_+ \times \mathbb{R}^n_+$ and $u_t(\cdot, \cdot, \omega) : \mathbb{R}^n_+ \times \mathbb{R}^n_+ \to \mathbb{R}$ is Lipschitz of rank $k_t(\omega)$ a.e. $\omega \in \Omega$, where $k_t : \Omega \to \mathbb{R}_+$ is an $\mathscr{F}_t$-measurable, integrable function.

(ii) There exists an $\mathscr{F}_t$-measurable, integrable function $\hat{u}_t : \Omega \to \mathbb{R}_+$ with $\sum_{t \in T} \int \hat{u}_t dP < +\infty$ satisfying $|u_t(z, z', \omega)| \le \hat{u}_t(\omega)$ for every $(z, z', \omega) \in G(Z_t)$.

Assumption 5.4. For every $x \in X_t$ and $y \in M_t(x)$: $(x, y) \in \operatorname{int} G(\Gamma_t)$.

Assumption 5.3 guarantees that f_t is Lipschitz of rank $\int k_t dP$ and Assumption 5.4 assures v_t is locally Lipschitz (see Lemma 4.2).

Lemma 5.2. *Suppose that Assumption 5.3 holds and let $(x, y) \in \operatorname{int} G(\Gamma_{t+1})$. If $u_{t+1}(\cdot, \cdot, \omega)$ is regular at $(x(\omega), y(\omega))$ for every $\omega \in \Omega$, then:*

(i) *f_t is regular at (x, y).*

(ii) *$\limsup_m f_t^\circ(x, y; h^m, k^m) \le f_t^\circ(x, y; h, k)$ whenever $(h^m, k^m) \to (h, k)$ a.e. in $L^\infty(\Omega, \mathscr{F}_t, P) \times L^\infty(\Omega, \mathscr{F}_{t+1}, P)$.*

Proof. (i): Take any $(h, k) \in L^\infty(\Omega, \mathscr{F}_t, P) \times L^\infty(\Omega, \mathscr{F}_{t+1}, P)$. For any $(z, z', \omega) \in G(Z_{t+1})$, $(a, a') \in \mathbb{R}^n \times \mathbb{R}^n$ and $\lambda > 0$, define:

$$\Delta u_{t+1}(z, z', \omega; a, a', \lambda) = u_{t+1}(z + \lambda a, z' + \lambda a', \omega) - u_{t+1}(z, z', \omega).$$

We then have:

$$\begin{aligned} f_t^\circ(x, y; h, k) &= \limsup_{\substack{(x', y') \to (x, y) \\ \lambda \downarrow 0}} \int_\Omega \frac{\Delta u_{t+1}(x'(\omega), y'(\omega), \omega; h(\omega), k(\omega), \lambda)}{\lambda} dP(\omega) \\ &\le \int_\Omega \limsup_{\substack{(z, z') \to (x(\omega), y(\omega)) \\ \lambda \downarrow 0}} \frac{\Delta u_{t+1}(z, z', \omega; h(\omega), k(\omega), \lambda)}{\lambda} dP(\omega) \end{aligned}$$

$$
\begin{aligned}
&= \lim_{\lambda \downarrow 0} \int_\Omega \frac{\Delta u_{t+1}(x(\omega), y(\omega), \omega; h(\omega), k(\omega), \lambda)}{\lambda} dP(\omega) \\
&= f'_t(x, y; h, k),
\end{aligned}
$$

where the first equality follows from the definition of the Clarke directional derivative, and the second inequality is a consequence of Fatou's lemma, noting that:

$$
\left| \frac{\Delta u_{t+1}(z, z', \omega; h(\omega), k(\omega), \lambda)}{\lambda} \right| \leq k_{t+1}(\omega)(\|h(\omega)\| + \|k(\omega)\|),
$$

for every $(z, z', \omega) \in G(Z_{t+1})$ and $\lambda > 0$ by Assumption 5.3(i), and that $(x', y') \to (x, y)$ in L^∞ implies convergence almost everywhere. The third equality uses the regularity of u_{t+1} and the Lebesgue-dominated convergence theorem, and the last equality follows from the definition of the directional derivative. Because the converse inequality $f'_t(x, y; h, k) \leq f^\circ_t(x, y; h, k)$ is always true, we obtain the regularity of f_t, as desired.

(ii): Let $(h^m, k^m) \to (h, k)$ a.e. in $L^\infty(\Omega, \mathscr{F}_t, P) \times L^\infty(\Omega, \mathscr{F}_{t+1}, P)$. We then have:

$$
\begin{aligned}
\limsup_{m \to \infty} f'_t(x, y; h^m, k^m) &= \limsup_{m \to \infty} \int_\Omega u'_{t+1}(x(\omega), y(\omega), \omega; h^m(\omega), k^m(\omega)) dP(\omega) \\
&\leq \int_\Omega \limsup_{m \to \infty} u'_{t+1}(x(\omega), y(\omega), \omega; h^m(\omega), k^m(\omega)) dP(\omega) \\
&= \int_\Omega u'_{t+1}(x(\omega), y(\omega), \omega; h(\omega), k(\omega)) dP(\omega) \\
&= f'_t(x, y; h, k),
\end{aligned}
$$

where the first equality uses the definition of the directional derivative and the Lebesgue-dominated convergence theorem, the second inequality exploits Fatou's lemma, the third equality employs the continuity of the directional derivative $(z, z') \mapsto u'_{t+1}(x(\omega), y(\omega), \omega; z, z')$, which is a consequence of the inequality $|u'_{t+1}(x(\omega), y(\omega), \omega; z, z')| \leq k_{t+1}(\omega)(\|z\| + \|z'\|)$ resulting from Assumption 5.3(i), and the final equality follows from the definition of the directional derivative and the Lebesgue-dominated convergence theorem. Because $f'_t(x, y; \cdot, \cdot) = f^\circ_t(x, y; \cdot, \cdot)$ by (i), we have $\limsup_m f^\circ_t(x, y; h^m, k^m) \leq f^\circ_t(x, y; h, k)$. □

Single-valuedness of the Directional Derivative

Because of Theorems 3.4 and 3.5, the single-valuedness of the directional derivative in Assumption 4.5 is satisfied if the convexity or supermodularity assumptions below are imposed on the primitive (u_t, Z_t) of the model.

Assumption 5.5. (i) $G(Z_t(\,\cdot\,,\omega))$ is a convex subset of $\mathbb{R}^n_+ \times \mathbb{R}^n_+$ a.e. $\omega \in \Omega$.

(ii) $u_t(\,\cdot\,,\cdot\,,\omega) : \mathbb{R}^n_+ \times \mathbb{R}^n_+ \to \mathbb{R}$ is concave a.e. $\omega \in \Omega$.

Assumption 5.5*. (i) $G(Z_t(\,\cdot\,,\omega))$ is a sublattice of $\mathbb{R}^n_+ \times \mathbb{R}^n_+$ a.e. $\omega \in \Omega$.

(ii) $u_t(\,\cdot\,,\cdot\,,\omega) : \mathbb{R}^n_+ \times \mathbb{R}^n_+ \to \mathbb{R}$ is supermodular a.e. $\omega \in \Omega$.

It is easy to see that Assumption 5.5 implies Assumption 4.4 and that Assumption 5.5* yields Assumption 4.6.

5.3 Price Systems in L^1

Let $ba(\Omega, \mathscr{F}_t)$ be the space of all $\mathbb{R}^n$-valued, bounded, finitely additive set functions (signed charges) on $\mathscr{F}_t$ and let $ba(\Omega, \mathscr{F}_t, P)$ be its vector subspace consisting of all elements that are absolutely continuous with respect to P. The dual space of $L^\infty(\Omega, \mathscr{F}_t, P)$ is $ba(\Omega, \mathscr{F}_t, P)$ (see Dunford and Schwartz [17, Theorem IV.8.16]). Let $ca(\Omega, \mathscr{F}_t)$ be the space of all $\mathbb{R}^n$-valued countably additive set functions (signed measures) on $\mathscr{F}_t$ and $ca(\Omega, \mathscr{F}_t, P)$ be its vector subspace consisting of all elements that are absolutely continuous with respect to P. A signed charge $\nu \in ba(\Omega, \mathscr{F}_t)$ is *purely finitely additive* if any $\mu \in ca(\Omega, \mathscr{F}_t)$ with $0 \le \mu \le |\nu|$ is identically zero, where $|\nu| = (|\nu_1|, \dots, |\nu_n|) \in ba(\Omega, \mathscr{F}_t)$ is the ($\mathbb{R}^n$-valued) total variation of $\nu = (\nu_1, \dots, \nu_n)$.

Proposition 5.2 (Yosida and Hewitt). (i) *Every $\nu \in ba(\Omega, \mathscr{F}_t)$ has a unique decomposition written as a sum $\nu = \nu^c + \nu^f$ such that ν^c is in $ca(\Omega, \mathscr{F}_t)$ and $\nu^f \in ba(\Omega, \mathscr{F}_t)$ is purely finitely additive.*

(ii) *A signed charge $\nu \in ba(\Omega, \mathscr{F}_t, P)$ is purely finitely additive if and only if there exists a sequence $\{A^m\}$ in $\mathscr{F}_t$ with $A^m \subset A^{m+1}$ for each $m = 1, 2, \dots$ satisfying $\lim_m P(A^m) = 1$ and $\langle \nu, x\chi_{A^m} \rangle = 0$ for every $x \in L^\infty(\Omega, \mathscr{F}_t, P)$ and $m = 1, 2, \dots$.*

This is a slightly generalized version of Yosida and Hewitt [39] in that the singularity of purely finitely additive parts in Proposition 5.2(ii) is allowed for signed charges in $ba(\Omega, \mathscr{F}_t, P)$. A proof of Proposition 5.2 can be found in Bhaskara Rao and Bhaskara Rao [10, Theorems 10.2.1 and 10.3.3].

Theorem 5.1. *Suppose that Assumptions 5.1 to 5.5, or Assumptions 5.1 to 5.4 and Assumption 5.5* are satisfied and let $(x_0, x_1, \dots) \in \mathscr{A}_0(x_0)$ be an optimal program. If $u_t(\,\cdot\,,\cdot,\,\omega)$ is regular at $(x_t(\omega), x_{t+1}(\omega))$ for every $\omega \in \Omega$, then there exists a price system $\{q_t \in L^1(\Omega, \mathscr{F}_t, P) \mid t \in T\}$ such that:*

(i) $q_t \in \partial v_t(x_t)$ *for every* $t \in T$.

(ii) $(q_t, -q_{t+1}) \in \partial f_t(x_t, x_{t+1})$ *for every* $t \in T$.

Proof. By Lemmas 5.1 and 5.2, the hypotheses in Theorem 4.2 are met. Let $\{p_t \in ba(\Omega, \mathscr{F}_t, P) \mid t \in T\}$ be a sequence in Theorem 4.2. It follows from Proposition 5.2 that p_t can be decomposed uniquely as $p_t = p_t^c + p_t^f$, where $p_t^c \in ca(\Omega, \mathscr{F}_t, P)$ and $p_t^f \in ba(\Omega, \mathscr{F}_t, P)$ is purely finitely additive, and there exists an increasing sequence $\{A^m\}$ in $\mathscr{F}_t$ such that $P(A^m) \to 1$ and $\langle p_t^f, x\chi_{A^m}\rangle = 0$ for every $x \in L^\infty(\Omega, \mathscr{F}_t, P)$ and $m = 1, 2, \dots$. Because $(p_t, -p_{t+1}) \in \partial f_t(x_t, x_{t+1})$ implies:

$$\begin{aligned} f_t^\circ(x_t, x_{t+1}; h\chi_{A^m}, k\chi_{A^m}) &\geq \langle p_t, h\chi_{A^m}\rangle - \langle p_{t+1}, k\chi_{A^m}\rangle \\ &= \langle p_t^c, h\chi_{A^m}\rangle - \langle p_{t+1}^c, k\chi_{A^m}\rangle \end{aligned}$$

for every $(h, k) \in L^\infty(\Omega, \mathscr{F}_t, P) \times L^\infty(\Omega, \mathscr{F}_{t+1}, P)$, we obtain:

$$\begin{aligned} f_t^\circ(x_t, x_{t+1}; h, k) &\geq \limsup_{m\to\infty} f_t^\circ(x_t, x_{t+1}; h\chi_{A^m}, k\chi_{A^m}) \\ &\geq \lim_{m\to\infty} [\langle p_t^c, h\chi_{A^m}\rangle - \langle p_{t+1}^c, k\chi_{A^m}\rangle] = \langle p_t^c, h\rangle - \langle p_{t+1}^c, k\rangle, \end{aligned}$$

in view of $(h\chi_{A^m}, k\chi_{A^m}) \to (h, k)$ a.e. in $L^\infty(\Omega, \mathscr{F}_t, P) \times L^\infty(\Omega, \mathscr{F}_{t+1}, P)$ and Lemma 5.2(ii). Therefore, $(p_t^c, -p_{t+1}^c) \in \partial f_t(x_t, x_{t+1})$. In this argument, putting $k = 0$ yields $f_{t,x}^\circ(x_t, x_{t+1}; h) \geq \langle p_t^c, h\rangle$ for every $h \in X$. Because Assumption 4.5 is satisfied by Assumption 5.5 or 5.5*, we have $\partial_x f_t(x_t, x_{t+1}) = \partial v_t(x_t)$ by Theorem 4.4. This is equivalent to $f_{t,x}^\circ(x_t, x_{t+1}; h) = v_t^\circ(x_t; h)$ for every $h \in X$ by (2.1). Thus, $p_t^c \in \partial v_t(x)$ follows from $v_t^\circ(x_t; h) \geq \langle p_t^c, h\rangle$ for every $h \in X$. Because p_t^c is absolutely continuous with respect to P, there exists a unique Radon–Nikodym derivative $q_t \in L^1(\Omega, \mathscr{F}_t, P)$ of p_t^c such that $p_t^c(A) = \int_A q_t dP$ for every $A \in \mathscr{F}_t$. Therefore, $p_t^c \in ca(\Omega, \mathscr{F}_t, P)$ can be identified with $q_t \in L^1(\Omega, \mathscr{F}_t, P)$ and the assertion holds. □

Unlike previous studies by Bewley [8, 9], Dana [14], Evstigneev [18], Jeanjean [21], Radner [31], Takekuma [36] and Zilcha [42, 43], we dispense with the monotonicity assumption on the random utility function and admit negative price systems in Theorem 5.1. To avoid the monotonicity assumption under the convexity hypothesis, Papageorgiou [28] applied a more general version of Proposition 5.2 in which bounded signed charges with Banach space values have Yosida–Hewitt decompositions (see Levin [23]).

6 Conclusion

While Theorem 5.1 assures the existence of L^1-price systems without convexity assumptions for the supermodular case with Assumption 5.5*, Assumption 5.2 is admittedly a more stringent condition than required in the convex model along the lines of Dana [14], Evstigneev [18], Jeanjean [21], Radner [31], Papageorgiou [28], Takekuma [36] and Zilcha [42, 43] for the derivation of an L^1-price system.

Indeed, there is an ambivalence between convexity and compactness. The nonconvex case under investigation requires only the Lipschitz continuity of f_t, which is automatically satisfied whenever u_t is continuous and concave, but we need the norm compactness of $\Gamma_t(x)$ to exploit subdifferential calculus for the nonconvex case in stochastic dynamic programming. In contrast, the concavity of u_t is needed in the convex case to ensure the weak*-continuity of f_t, but the weak*-compactness of $\Gamma_t(x)$ is sufficient to guarantee the existence of optimal programs, and is easily satisfied under the boundedness condition on Z_t. The usual argument employing the separation theorem is valid for detecting a support price system.

Because support prices are unavailable in the nonconvex case, the traditional approach does not work for the derivation of the price systems in L^1. To overcome this difficulty, this paper focuses on the necessary condition for optimality in Theorem 4.2 described in terms of the Clarke subdifferential. In particular, when f_t is concave, the Clarke directional derivative in the proof of Theorem 5.1 reduces to the directional derivative; thus our approach is essentially the same as the traditional one except for the topological setting. However, exploiting the Clarke directional derivative in the limiting argument presents a more general method that likewise works for the nonconvex case.

References

[1] Aliprantis, C. D. and K. C. Border, (2006). *Infinite Dimensional Analysis: A Hitchhiker's Guide*, 3rd edn., Berlin, Springer-Verlag.

[2] Aliprantis, C. D., G. Camera and F. Ruscitti, (2009). "Monetary equilibrium and the differentiability of the value function", *J. Econom. Dynam. Control* **33**, 454–462.

[3] Amir, R., (1996). "Sensitivity analysis of multisector optimal economic dynamics", *J. Math. Econom.* **25**, 123–141.

[4] Askri, K. and C. Le Van, (1998). "Differentiability of the value function of nonclassical optimal growth models", *J. Optim. Theory Appl.* **97**, 591–604.

[5] Aubin, J.-P. and I. Ekeland, (1984). *Applied Nonlinear Analysis*, New York, John Wiley & Sons.

[6] Aumann, R. J., (1967). "Measurable utility and the measurable choice theorem", *La décision*, Actes Coll. Int. du CNRS, Aix-en-Provence, Paris, 15–26.

[7] Benveniste, L. M. and J. A. Scheinkman, (1979). "On the differentiability of the value function in dynamic models of economics", *Econometrica* **47**, 727–732.

[8] Bewley, T. F., (1972). "Existence of equilibria in economies with infinitely many commodities", *J. Econom. Theory* **4**, 514–540.

[9] Bewley, T. F., (1981). "Stationary equilibrium", *J. Econom. Theory* **24**, 265–295.

[10] Bhaskara Rao, K. P. S. and M. Bhaskara Rao, (1983). *Theory of Charges: A Study of Finitely Additive Measures*, London, Academic Press.

[11] Bonnisseau, J.-M. and C. Le Van, (1996). "On the subdifferential of the value function in economic optimization problems", *J. Math. Econom.* **25**, 55–73.

[12] Borwein, J. M. and Q. J. Zhu, (2005). *Techniques of Variational Analysis*, Berlin, Springer-Verlag.

[13] Clarke, F. H., (1983). *Optimization and Nonsmooth Analysis*, New York, John Wiley & Sons.

[14] Dana, R.-A., (1974). "Evaluation of development programs in a stationary stochastic economy with bounded primary resources", in: J. Løs and M. W. Løs, (eds.), *Mathematical Models in Economics*, Amsterdam, North-Holland, pp. 179–205.

[15] Danskin, J. M., (1967). *The Theory of Max Min*, Berlin, Springer-Verlag.

[16] Dontchev, A. and T. Zolezzi, (1993). *Well-Posed Optimization Problems*, Berlin, Springer-Verlag.

[17] Dunford, N. and J. T. Schwartz, (1958). *Linear Operators, Part I: General Theory*, New York, John Wiley & Sons.

[18] Evstigneev, I., (1974). "Optimal stochastic programs and their stimulating prices", in: J. Løs and M. W. Løs, (eds.), *Mathematical Models in Economics*, Amsterdam, North-Holland, pp. 219–252.

[19] Gale, D., (1967). "On optimal development in a multi-sector economy", *Rev. Econom. Stud.* **34**, 1–18.

[20] Ioffe, A. D. and V. L. Levin, (1972). "Subdifferentials of convex functions", *Trans. Moscow Math. Soc.* **26**, 1–72.

[21] Jeanjean, P., (1974). "Optimal development programs under uncertainty: The discounted case", *J. Econom. Theory* **7**, 66–92.

[22] Kim, T., (1993). "Differentiability of the value function: A new characterization", *Seoul J. Econom.* **6**, 257–265.

[23] Levin, V. L., (1974). "The Lebesgue decomposition for functionals on the vector function space L_X^∞", *Funct. Anal. Appl.* **8**, 314–317.

[24] McKenzie, L. W., (1986). "Optimal economic growth: Turnpike theorems and comparative dynamics", in: K. J. Arrow and M. D. Intriligator, (eds.), *Handbook of Mathematical Economics, Volume III*, Amsterdam, North-Holland, pp. 1281–1355.

[25] Michel, P., (1990). "Some clarifications on the transversality condition", *Econometrica* **58**, 705–723.

[26] Milgrom, P. and I. Segal, (2002). "Envelope theorems for arbitrary choice set", *Econometrica* **70**, 583–601.

[27] Mordukhovich, B. S., (2006). *Variational Analysis and Generalized Differentiation I: Basic Theory; II: Applications*, Berlin, Springer-Verlag.

[28] Papageorgiou, N. S., (1994). "Optimal programs and their price characterization in a multisector growth model with uncertainty", *Proc. Amer. Math. Soc.* **122**, 227–240.

[29] Penot, J.-P., (2004). "Differentiability properties of optimal value functions", *Canad. J. Math.* **56**, 825–842.

[30] Pshenichnyi, B. N., (1971). *Necessary Conditions for an Extremal*, New York, Marcel Dekker.

[31] Radner, R., (1973), "Optimal stationary consumption with stochastic production and resources", *J. Econom. Theory* **6**, 68–90.

[32] Rincón-Zapatero, J. P. and M. S. Santos, (2009). "Differentiability of the value function without interiority assumptions", *J. Econom. Theory* **144**, 1948–1964.

[33] Sagara, N., (2010). "Value functions and transversality conditions for infinite-horizon optimal control problems", *Set-Valued Var. Anal.* **18**, 1–28.

[34] Schirotzek, W., (2007). *Nonsmooth Analysis*, Berlin, Springer-Verlag.

[35] Stokey, N. L. and R. E. Lucas, Jr., (1989). *Recursive Methods in Economic Dynamics*, Cambridge, Massachusetts, Harvard University Press.

[36] Takekuma, S.-I., (1992). "Optimal growth under uncertainty: A complete characterization of weakly maximal programs", *Hitotsubashi J. Econom.* **33**, 169–182.

[37] Thibault, L., (1991). "On subdifferentials of optimal value functions", *SIAM J. Control Optim.* **29**, 1019–1036.

[38] Topkis, D. M., (1978). "Minimizing a submodular function on a lattice", *Manag. Sci.* **26**, 305–321.

[39] Yosida, K. and E. Hewitt, (1952). "Finitely additive measures", *Trans. Amer. Math. Soc.* **72**, 46–66.

[40] Weitzman, M. L., (1973). "Duality theory for infinite horizon convex models", *Manag. Sci.* **19**, 783–789.

[41] Zălinescu, C., (2002). *Convex Analysis in General Vector Spaces*, Singapore, World Scientific.

[42] Zilcha, I., (1976). "On competitive prices in a multi-sector economy with stochastic production and resources", *Rev. Econom. Stud.* **43**, 431–438.

[43] Zilcha, I., (1976). "Characterization by prices of optimal programs under uncertainty", *J. Math. Econom.* **3**, 173–183.